MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2019/2020

BEC2034 – BASIC ECONOMETRICS

(All Sections/ Groups)

21 OCTOBER 2019
2.30 p.m. - 4.30 p.m.
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This Question paper consists of **EIGHT** pages with **FOUR** questions only.
2. Attempt **ALL FOUR** questions. The distribution of the marks for each question is given.
3. Please write all your answers in the Answer Booklet provided.
4. Formulas and statistical tables are attached in pages 6–8.

QUESTION 1 (25 marks)

(a) A researcher estimates the following two econometric models:

$$y_t=\beta_1+\beta_2x_{2t}+\beta_3x_{3t}+u_t \quad \ldots (1.1)$$
$$y_t=\beta_1+\beta_2x_{2t}+\beta_3x_{3t}+\beta_4x_{4t}+v_t \quad \ldots (1.2)$$

where u_t and v_t are independent and identically distributed (iid) disturbances and x_{4t} is an irrelevant variable which does not enter into the data generating process for y_t. Will the following value be higher in the model (1.2) than in the model (1.1)? Explain your answers.

(i) R^2 (5 marks)
(ii) Adjusted R^2 (5 marks)

(b) "Heteroscedasticity, when left unchecked, can leave a research which has found statistically significant results open to criticism". Explain. (5 marks)

(c) The researcher estimates a model with the dependent variable of the monthly returns on a given share traded on the New York Stock Exchange, and various macroeconomic variables and financial ratios as independent variables. She attempts to estimate this model, together with twelve monthly dummy variables (January to December, one for each month), and a constant term, using EViews. The result of the EViews shows that the parameters of the model cannot be estimated. Provide **TWO (2)** solutions to how the researcher can fix the problem. (10 marks)

QUESTION 2 (25 marks)

A research has been done to determine the value of houses in Kajang. The estimated model is as follows:

$$\text{VALUE}_i = \alpha_1 + \alpha_2\text{CRIME}_i + \alpha_3\text{NITOX}_i + \alpha_4\text{ROOMS}_i + \alpha_5\text{DIST}_i + \alpha_6\text{ACCESS}_i + \alpha_7\text{PTRATIO}_i + e_i \quad \ldots (2.1)$$

where

VALUE	=	medium value of owner-occupied homes in RM'000;
CRIME	=	per capita crime rate;
NITOX	=	nitric oxide concentration (parts per million);
ROOMS	=	average number of rooms per dwelling;
DIST	=	weighted distance to five Kajang employment centres;
ACCESS	=	index of accessibility to radial highways;
PTRATIO	=	pupil-teacher ratio by town.

Continued...

The output of regression model **(2.1)** is shown below:

Dependent Variable: VALUE
Include observations: 506

Variable	Beta	Standard Error
Constant	28.4067	5.3659
CRIME	-0.1834	0.0365
NITOX	-22.8109	4.1607
ROOMS	6.3715	0.3924
DIST	-1.3353	0.2001
ACCESS	0.2723	0.0723
PTRATIO	-1.1768	0.1394

$R^2 = 0.657$ Jarque-Bera statistics = 4.99 White statistics = 42.84

(a) Report briefly on how each of the variables influences the value of a home. (3 marks)

(b) Compute the value of adjusted R^2. (2 marks)

(c) Is the error term normally distributed? Use 5% as the level of significance. (5 marks)

(d) Is the overall regression significant? Use 5% as the level of significance. (5 marks)

(e) Test as an alternative hypothesis H_1 that there is a negative relationship between the pupil-teacher ratio and the value of a house. Use 5% as the level of significance. (5 marks)

(f) Test the hypothesis that increasing the number of rooms by one, the value of a house will increase by RM7,000. Use 5% as the level of significance. (5 marks)

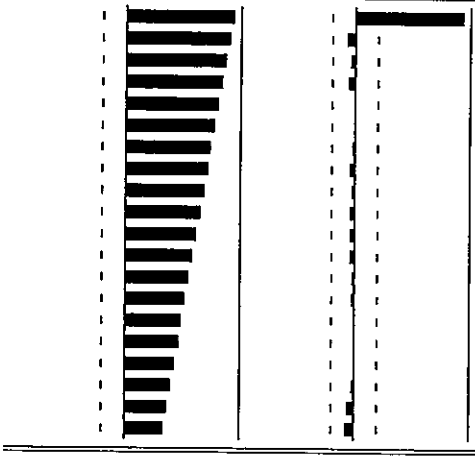
Continued...

QUESTION 3 (25 marks)

(a) What is the meaning of a unit root? (6 marks)

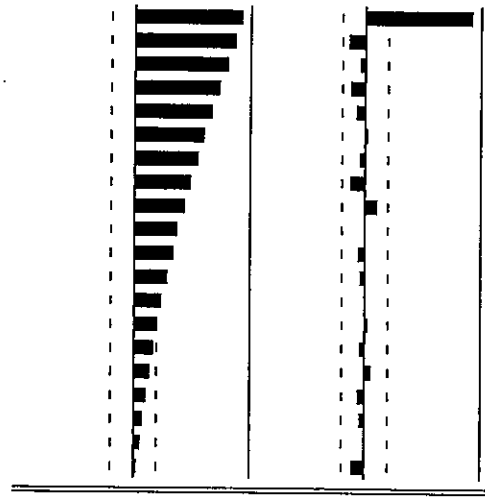
(b) The Sample Correlograms (up to 20 lags) for the time series of corporate financial performance (CFP) and corporate social responsibility (CSR) are shown as below:

Time series: CFP



	AC	PAC	Q-Stat	Prob
1	0.972	0.972	85.968	0.000
2	0.941	-0.063	167.49	0.000
3	0.909	-0.035	244.44	0.000
4	0.875	-0.043	316.64	0.000
5	0.842	-0.005	384.29	0.000
6	0.810	-0.005	447.61	0.000
7	0.777	-0.018	506.70	0.000
8	0.744	-0.035	561.52	0.000
9	0.711	-0.016	612.21	0.000
10	0.677	-0.027	658.82	0.000
11	0.643	-0.027	701.40	0.000
12	0.609	-0.025	740.07	0.000
13	0.575	-0.013	775.01	0.000
14	0.542	-0.008	806.46	0.000
15	0.510	0.001	834.70	0.000
16	0.480	0.004	860.04	0.000
17	0.451	-0.001	882.72	0.000
18	0.422	-0.022	902.87	0.000
19	0.392	-0.044	920.49	0.000
20	0.359	-0.076	935.46	0.000

Time Series: CSR



	AC	PAC	Q-Stat	Prob
1	0.954	0.954	82.833	0.000
2	0.897	-0.145	156.90	0.000
3	0.838	-0.031	222.39	0.000
4	0.772	-0.123	278.55	0.000
5	0.701	-0.061	325.49	0.000
6	0.636	0.023	364.55	0.000
7	0.572	-0.027	396.56	0.000
8	0.501	-0.119	421.45	0.000
9	0.443	0.102	441.10	0.000
10	0.391	0.004	456.65	0.000
11	0.339	-0.047	468.47	0.000
12	0.289	-0.026	477.17	0.000
13	0.245	-0.001	483.51	0.000
14	0.206	0.019	488.06	0.000
15	0.168	-0.025	491.12	0.000
16	0.139	0.042	493.25	0.000
17	0.111	-0.049	494.63	0.000
18	0.081	-0.032	495.38	0.000
19	0.055	-0.002	495.72	0.000
20	0.024	-0.104	495.79	0.000

What general pattern do you see? Intuitively, do both, either, neither of these time series seem to be stationary? Explain. (4 marks)

Continued...

(c) The Dickey-Fuller (DF) test results for these two time series are as below:

$$\begin{aligned} \Delta CSR_t &= 6.522 + 0.084t - 0.069CSR_{t-1} \\ \tau &= (2.154) \quad (1.142) \quad (-1.715) \end{aligned} \qquad R^2 = 0.037 \quad \ldots (3.1)$$

$$\begin{aligned} \Delta CFP_t &= 0.565 + 0.113t - 0.063CFP_{t-1} \\ \tau &= (1.515) \quad (3.138) \quad (-2.640) \end{aligned} \qquad R^2 = 0.148 \quad \ldots (3.2)$$

The critical 5 percent τ value = –3.46.

Find out if each of these series contains a unit root. If a unit root exists, how would you characterise such a time series? (5 marks)

(d) Since corporate financial performance (CFP) depends on corporate social responsibility (CSR), consider the following simple model:

$$\begin{aligned} CFP &= -13.3114 + 0.6281\ CSR \\ t &\quad (-1.8079) \quad (11.9253) \end{aligned} \qquad R^2 = 0.6231 \quad d = 0.0712 \quad \ldots (3.3)$$

(i) Based on the test results in (c) and the equation (3.3), would you expect this regression to suffer from the spurious regression phenomenon? Why? (5 marks)

(ii) How do you explicitly test whether CFP and CSR time series are cointegrated? Elaborate your answer in details (5 marks)

Continued...

QUESTION 4 (25 marks)

(a) Discuss **THREE** advantages of panel data over cross-section or time-series data. (12 marks)

(b) Based on the Michigan Income Dynamics Study, Hausman attempted to estimate a wage, or earnings, model using a sample of 629 high school graduates, who were followed for a period of 6 years, thus giving in all 3774 observations. The dependent variable in this study was logarithm of wage (Log Wage), and the explanatory variables were age (divided into several age groups), unemployment in the previous year, poor health in the previous year, self-employment, region of residence (South = 1; 0 otherwise), area of residence (Rural = 1; 0 otherwise). Hausman used both fixed effects model (FEM) and error components model (ECM). The results are given in **Table 4.1** (standard errors in parentheses):

Table 4.1 Wage Equations (Dependent Variable: Log Wage*)

	Fixed effects		**Error Components**	
Variable	**Estimated Coefficient**	**Standard error**	**Estimated Coefficient**	**Standard error**
Age 1 (20-34)	0.0557	(0.0042)	0.0393	(0.0033)
Age 2 (35-44)	0.0351	(0.0051)	0.0092	(0.0036)
Age 3 (45-54)	0.0209	(0.0055)	-0.0007	(0.0042)
Age 4 (55-64)	0.0209	(0.0078)	-0.0097	(0.0060)
Age 5 (65 and above)	-0.0171	(0.0155)	-0.0423	(0.0121)
Unemployed previous year	-0.0042	(0.0153)	-0.0277	(0.0151)
Poor health previous year	-0.0204	(0.0221)	-0.0250	(0.0215)
Self-employment	-0.2190	(0.0297)	-0.2670	(0.0263)
South	-0.1599	(0.0656)	-0.0324	(0.0333)
Rural	-0.0101	(0.0317)	-0.1215	(0.0237)
Constant	-		0.8499	(0.0433)
Degree of freedom	3,135		3,763	

Note: *3774 observations.
Reproduced from Cheng Hsiao, Analysis of Panel Data, Cambridge University Press, 1986, p.42. Original source: J.A. Hausman, "Specification Tests in Econometrics," Econometrica, vol 46, 1978, pp. 1251-1271.

(i) Based on the estimated coefficients for both model specifications, do the results make economic sense? (4 marks)

(ii) On the basis of the data given in the **Table 4.1**, which model, if any, would you choose? Briefly explain your selection. (3 marks)

(iii) State the hypothesis and explain the steps involved in choosing fixed effects or error component model using the Hausman test. (6 marks)

End of Questions

Formulas

$$\text{t - statistic} = \frac{\hat{\beta} - \beta}{\text{se}(\hat{\beta})}$$

$$R^2 = \frac{\text{Explained Sum of Squares}}{\text{Total Sum of Squares}}$$

$$\text{Adjusted } R^2 = 1 - (1 - R^2)\left(\frac{n-1}{n-k-1}\right)$$

Total Sum of Squares (TSS) = Explained Sum of Squares (ESS) + Residual Sum of Squares (RSS)

$$\text{F - statistic} = \frac{\text{ESS}/(k)}{\text{RSS}/(n-k-1)} = \frac{R^2/(k)}{(1-R^2)/(n-k-1)}$$

White's general heteroscedasticity test statistics, $\chi^2_{df=k} = n \times R^2$

rho, $\rho \approx 1 - (d/2)$

Tolerance = $(1 - R^2_{\text{auxiliary regression}})$

Statistical Tables

Appendix 1: t-Table

two tails	**0.2**	**0.1**	**0.05**	**0.02**	**0.01**
One tail	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**
df					
10	1.37	1.81	2.23	2.76	3.17
20	1.33	1.72	2.09	2.53	2.84
30	1.31	1.70	2.04	2.46	2.75
40	1.30	1.68	2.02	2.42	2.70
50	1.30	1.68	2.01	2.40	2.68
60	1.30	1.67	2.00	2.39	2.66
75	1.29	1.67	1.99	2.38	2.64
100	1.29	1.66	1.98	2.36	2.63
120	1.29	1.66	1.98	2.36	2.62

Appendix 2: F-table (α=0.05)

df2\df1	1	2	3	4	5	6	7	8
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06
100	3.94	3.09	2.7	2.46	2.31	2.19	2.10	2.03
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96
1000	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95

Appendix 3: Chi-square table (α=0.05)

df	α = 0.05
1	3.84
2	5.99
3	7.82
4	9.49
5	11.07
10	18.31
20	31.41
30	43.77
40	55.76
50	67.51
100	124.34

Appendix 4: Durbin Watson d Table

Critical Values for the Durbin-Watson Statistic (d)

Level of Significance $\alpha = 0.05$

n	$k=1$		$k=2$		$k=3$		$k=4$		$k=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
10	0.88	1.32	0.70	1.64	0.53	2.02	0.38	2.41	0.24	2.82
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
25	1.29	1.41	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78
150	1.72	1.75	1.71	1.76	1.69	1.77	1.68	1.79	1.66	1.80
200	1.76	1.78	1.75	1.79	1.74	1.80	1.73	1.81	1.72	1.82

n	$k=6$		$k=7$		$k=8$		$k=9$		$k=10$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
10	–	–	–	–	–	–	–	–	–	–
15	0.45	2.47	0.34	2.73	0.25	2.98	0.18	3.22	0.11	3.44
20	0.69	2.16	0.60	2.34	0.50	2.52	0.42	2.70	0.34	2.89
25	0.87	2.01	0.78	2.14	0.70	2.28	0.62	2.42	0.54	2.56
30	1.00	1.93	0.93	2.03	0.85	2.14	0.78	2.25	0.71	2.36
40	1.18	1.85	1.12	1.92	1.06	2.00	1.01	2.07	0.95	2.15
50	1.29	1.82	1.25	1.88	1.20	1.93	1.16	1.99	1.11	2.04
60	1.37	1.81	1.34	1.85	1.30	1.89	1.26	1.94	1.22	1.98
70	1.43	1.80	1.40	1.84	1.37	1.87	1.34	1.91	1.31	1.95
80	1.48	1.80	1.45	1.83	1.43	1.86	1.40	1.89	1.37	1.93
90	1.52	1.80	1.49	1.83	1.47	1.85	1.45	1.88	1.42	1.91
100	1.55	1.80	1.53	1.83	1.50	1.85	1.48	1.87	1.46	1.90
150	1.65	1.82	1.64	1.83	1.62	1.85	1.60	1.86	1.59	1.88
200	1.71	1.83	1.70	1.84	1.69	1.85	1.68	1.86	1.67	1.87

where n = number of observations and k = number of independent variables, excluding the intercept.

End of Page